

Review of LFTs, LMIs, and μ

John Doyle[†]

Andy Packard[‡]

Kemin Zhou[§]

Abstract

The purpose of this paper is to present a tutorial overview of Linear Fractional Transformations (LFT) and the role of the Structured Singular Value, μ , and Linear Matrix Inequalities (LMI) in solving LFT problems.

1 Introduction

LFTs and LMIs play a very important role in postmodern control theory by providing a framework that unifies many concepts and generalizes transfer functions and their state-space realizations to include uncertainty. The focus of this paper is on reviewing known results on robust stability and performance and establishing a common and unified framework for the companion papers in this session, which consider generalizations and extensions of balanced realizations and model reduction [WDBG], stabilization [LuZD], optimal control [PZPB], mixed real/complex μ [YoND], model validation [New], and LMI computation [Beck].

Section 2 introduces the notation for LFTs and briefly discusses some of their properties. Section 3 describes μ and its connections with LFTs. Section 4 focuses on two standard notions of robust stability and performance, μ stability and performance and Q stability and performance, and their relationship is discussed. Comparisons with the new and exciting L_1 theory of robust performance with structured uncertainty are also considered.

2 Linear Fractional Transformations (LFTs)

2.1 Definitions of LFTs

Suppose M is a complex matrix partitioned as

$$M = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} \in \mathbb{C}^{(p_1+p_2) \times (q_1+q_2)} \quad (2.1)$$

and let $D_1 \subset \mathbb{C}^{q_1 \times p_1}$ and $D_2 \subset \mathbb{C}^{q_2 \times p_2}$, then we define the linear fractional transformations (LFTs) as the maps:

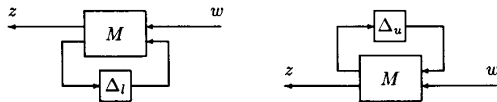
$$\mathcal{F}_l(M, \bullet) : D_2 \mapsto \mathbb{C}^{p_1 \times q_1} \quad \mathcal{F}_u(M, \bullet) : D_1 \mapsto \mathbb{C}^{p_2 \times q_2}$$

with

$$\mathcal{F}_l(M, \Delta_l) := M_{11} + M_{12}\Delta_l(I - M_{22}\Delta_l)^{-1}M_{21} \quad (2.2)$$

$$\mathcal{F}_u(M, \Delta_u) := M_{22} + M_{21}\Delta_u(I - M_{11}\Delta_u)^{-1}M_{12} \quad (2.3)$$

Clearly the existence of the inverses is necessary for the LFT's to be well defined. We can also define the LFTs more generally, say with respect to a real rational matrix $\Delta \in \mathbb{R}^{n \times m}(s)$, with the other related matrices also being defined as real rational. The LFT formulae arise naturally when describing feedback systems as shown in the following figures.



^{*}Edited by John Doyle from material by Andy Packard and Kemin Zhou. With help from Peter Young, Carolyn Beck, Jorge Tierno, and Way Lu, and support from NSF, ONR, NASA, and AFOSR.

[†]Electrical Engineering, M/S 116-81, Caltech, Pasadena, CA 91125

[‡]Mechanical Engineering, UC Berkeley, Berkeley, CA 94720

[§]Electrical and Computer Engineering, LSU, Baton Rouge, LA 70803

The resulting closed-loop transfer functions from w to z are, respectively, $\mathcal{F}_l(M, \Delta_l)$ and $\mathcal{F}_u(M, \Delta_u)$.

2.2 Redheffer Star-Products

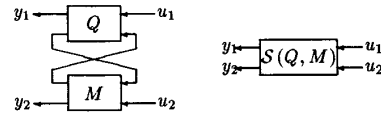
Suppose that Q and M are complex matrices, partitioned as

$$Q = \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix}, \quad M = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix}$$

with the matrix product $Q_{22}M_{11}$ well defined, and in fact, square. If $I - Q_{22}M_{11}$ is invertible, define the star product of Q and M , with respect to this partition to be

$$S(Q, M) := \begin{bmatrix} \mathcal{F}_l(Q, M_{11}) & Q_{12}(I - M_{11}Q_{22})^{-1}M_{12} \\ M_{21}(I - Q_{22}M_{11})^{-1}Q_{21} & \mathcal{F}_u(M, Q_{22}) \end{bmatrix}$$

Note that this definition is dependent on the partitioning of the matrices Q and M above. In fact it may be well defined for one partition and not well defined for another. However, we will not explicitly show this dependence, as it is always clear from the context. In a block diagram, this appears as



Recall the definitions of linear fractional transformations. Note that for a matrix K of the appropriate dimension, if all of the necessary matrices are invertible (implied by the loop equations) then

$$\mathcal{F}_l(S(Q, M), K) = \mathcal{F}_l(Q, \mathcal{F}_l(M, K))$$

We can also use the S notation for LFTs, as in

$$S(M, \Delta_l) = \mathcal{F}_l(M, \Delta_l) \quad S(\Delta_u, M) = \mathcal{F}_u(M, \Delta_u)$$

2.3 Examples of LFTs

State Space, Transfer Functions and LFTs

Given the state space realization of a discrete time system

$$\begin{bmatrix} x_{k+1} \\ y_k \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} x_k \\ u_k \end{bmatrix} = M \begin{bmatrix} x_k \\ u_k \end{bmatrix} \quad (2.4)$$

then its transfer matrix is

$$G(z) = D + C(zI - A)^{-1}B = \mathcal{F}_u \left(\begin{bmatrix} A & B \\ C & D \end{bmatrix}, \frac{1}{z}I \right) =: \begin{bmatrix} A & B \\ C & D \end{bmatrix}$$

This last notation is deliberately somewhat ambiguous, and can be viewed as both a transfer matrix and its realization. The ambiguity is benign and convenient and can always be resolved from the context. We also use this notation for arbitrary LFTs when the arguments are clear from context, for example,

$$\mathcal{F}_u(M, \Delta_u) = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix}$$

Frequency Transformation

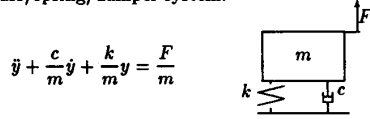
The bilinear transformation between the z -domain and s -domain is

$$s = \frac{z+1}{z-1}, \quad \frac{1}{s}I = I - \sqrt{2}I \quad z^{-1}I (I + z^{-1}I)^{-1} \sqrt{2}I = \mathcal{F}_u(N, z^{-1}I)$$

$$\text{where } N = \begin{bmatrix} I & \sqrt{2}I \\ -\sqrt{2}I & -I \end{bmatrix}$$

State Space Parametric Uncertainty

One natural type of uncertainty is unknown coefficients in a state space model. As a simple example, we will begin with a familiar idealized mass/spring/damper system.



$$\ddot{y} + \frac{c}{m}\dot{y} + \frac{k}{m}y = \frac{F}{m}$$

Suppose $m, c,$ and k are fixed but uncertain, with $m = \bar{m}(1 + w_m \delta_m), c = \bar{c}(1 + w_c \delta_c), k = \bar{k}(1 + w_k \delta_k)$. Then defining $x_1 = y$ and $x_2 = m\dot{y}$ we can write the differential equation in state-space form as

$$\begin{bmatrix} \dot{x} \\ y \end{bmatrix} = \mathcal{F}_l(M, \Delta) \begin{bmatrix} x \\ F \end{bmatrix} \quad \Delta = \text{diag}(\delta_m, \delta_k, \delta_c)$$

$$M = \begin{bmatrix} 0 & \bar{m}^{-1} & 0 & -w_m & 0 & 0 \\ -\bar{k} & -\bar{c} & 1 & 0 & w_k & w_c \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & \bar{m}^{-1} & 0 & -w_m & 0 & 0 \\ -\bar{k} & 0 & 0 & 0 & 0 & 0 \\ 0 & -\bar{c} & 0 & 0 & 0 & 0 \end{bmatrix}$$

More generally, the perturbed state-space system

$$\begin{aligned} x_{k+1} &= A(\delta)x_k + B(\delta)d_k \\ e_k &= C(\delta)x_k + D(\delta)d_k \end{aligned}$$

where δ is a vector of parameters that enter rationally can be written as an LFT on a diagonal matrix Δ made up of the elements of δ , possibly repeated. The form of the LFT is ([MorM])

$$\begin{bmatrix} x_{k+1} \\ e_k \\ z_k \end{bmatrix} = \begin{bmatrix} M_{11} & M_{12} & M_{13} \\ M_{21} & M_{22} & M_{23} \\ M_{31} & M_{32} & M_{33} \end{bmatrix} \begin{bmatrix} x_k \\ d_k \\ w_k \end{bmatrix} \quad (2.5)$$

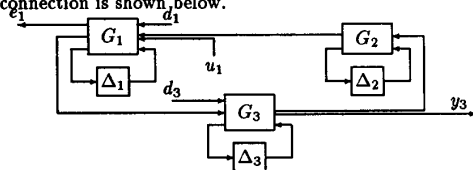
with perturbation $w_k = \Delta z_k$ yielding

$$\begin{bmatrix} x_{k+1} \\ e_k \end{bmatrix} = \mathcal{F}_l(M, \Delta) \begin{bmatrix} x_k \\ d_k \end{bmatrix} \quad (2.6)$$

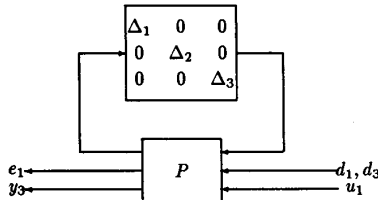
In general, for problems of this type it is easy to obtain realizations, but it is difficult to insure that they are minimal, except in the case where the parameters enter linearly.

Interconnections

Interconnections of LFTs are again LFTs. This is a fundamental property of linear fractional transformations, and is one reason why they are so important in linear systems theory. For example, consider a situation with three components, each with a LFT uncertainty model. The interconnection is shown below.



By simply reorganizing the diagram, collecting all of the known systems together, and collecting all of the perturbations (the Δ_i 's) together, we end up with the following diagram



Note how general uncertainty at the component level becomes structured uncertainty at the system level.

2.4 Properties of LFTs

One of the features of LFTs is that they can be manipulated much like state-space realizations of transfer functions. We can cascade, add, invert them and so on. Some examples are given below.

Operations On LFTs

Given two systems with realizations

$$G_1(\Delta_1) = \begin{bmatrix} A_1 & B_1 \\ C_1 & D_1 \end{bmatrix} \quad G_2(\Delta_2) = \begin{bmatrix} A_2 & B_2 \\ C_2 & D_2 \end{bmatrix}$$

define $\Delta = \begin{bmatrix} \Delta_1 & 0 \\ 0 & \Delta_2 \end{bmatrix}$. Then the cascade system has a realization

$$(G_1 G_2)(\Delta) = \begin{bmatrix} A_1 & B_1 C_2 & B_1 D_2 \\ 0 & A_2 & B_2 \\ C_1 & D_1 C_2 & D_1 D_2 \end{bmatrix}$$

and the addition of G_1 and G_2 has a realization

$$(G_1 + G_2)(\Delta) = \begin{bmatrix} A_1 & 0 & B_1 \\ 0 & A_2 & B_2 \\ C_1 & C_2 & D_1 + D_2 \end{bmatrix}$$

Inversion formulas

Suppose $\mathcal{F}_l(M, \Delta)$ is square and well-defined for all desired Δ and M_{11} is nonsingular. Then $(\mathcal{F}_l(M, \Delta))^{-1} = \mathcal{F}_l(\tilde{M}, \Delta)$ with \tilde{M} given by

$$\tilde{M} = \begin{bmatrix} M_{11}^{-1} & -M_{11}^{-1}M_{12} \\ M_{21}M_{11}^{-1} & M_{22} - M_{21}M_{11}^{-1}M_{12} \end{bmatrix}$$

Suppose that $G = \mathcal{F}_l(P, K)$ with $P, P_{12},$ and P_{21} are all square and nonsingular. Then we can solve for K and $K = \mathcal{F}_u(P^{-1}, G)$. This formula is easily verified by writing the equations for the LFT

$$\begin{bmatrix} z \\ y \end{bmatrix} = P \begin{bmatrix} w \\ u \end{bmatrix}, \quad u = Ky \\ z = \mathcal{F}_l(P, K)w = Gw$$

and solving them to yield

$$\begin{bmatrix} w \\ u \end{bmatrix} = P^{-1} \begin{bmatrix} z \\ y \end{bmatrix}, \quad z = Gw \\ u = \mathcal{F}_u(P^{-1}, G)y \\ K = \mathcal{F}_u(P^{-1}, G)$$

3 Structured Singular Value

3.1 Definitions

We consider matrices $M \in \mathbb{C}^{n \times n}$ and an underlying block structure Δ , (a prescribed set of block diagonal matrices) on which everything in the sequel depends. In this paper we will only consider the purely complex case (i.e. the block structure contains only complex uncertainties). For the mixed real and complex case see [YoND].

Two nonnegative integers, S and F , represent the number of repeated scalar blocks and the number of full blocks, respectively.

$$\Delta = \left\{ \text{diag} [\delta_1 I_{k_1}, \dots, \delta_s I_{k_s}, \Delta_1, \dots, \Delta_F] : \delta_i \in \mathbb{C}, \Delta_i \in \mathbb{C}^{k_{s+i} \times k_{s+i}} \right\} \quad (3.7)$$

$$\mathbf{B}\Delta = \{\Delta \in \Delta : \bar{\sigma}(\Delta) \leq 1\} \quad (3.8)$$

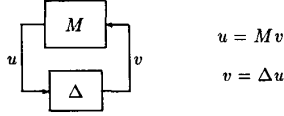
For notational convenience all of the repeated scalar blocks appear first, and the full blocks are square, but this is easily relaxed.

Definition 3.1 For $M \in \mathbb{C}^{n \times n}$, $\mu_\Delta(M)$ is defined

$$\mu_\Delta(M) := \frac{1}{\min \{\bar{\sigma}(\Delta) : \Delta \in \Delta, \det(I - M\Delta) = 0\}} \quad (3.9)$$

unless no $\Delta \in \Delta$ makes $I - M\Delta$ singular, in which case $\mu_\Delta(M) := 0$.

It follows immediately that $\mu_\Delta(M) = \max_{\Delta \in \mathbf{B}\Delta} \rho(M\Delta)$ where ρ denotes the spectral radius. It is instructive to consider the following “feedback” interpretation of $\mu_\Delta(M)$.



If $I - M\Delta$ is nonsingular, then the only solutions are $u = v = 0$, while if $I - M\Delta$ is singular, then there are infinitely many solutions and the norms $\|u\|, \|v\|$ of the solutions can be arbitrarily large. With a slight abuse of convention, we might call these systems “stable” and “unstable,” respectively. Viewed this way, $\mu_\Delta(M)$ is a measure of the smallest structured Δ that causes “instability.”

We can relate $\mu_\Delta(M)$ to familiar linear algebra quantities when Δ is one of two extreme sets.

- If $\Delta = \{\delta I : \delta \in \mathbb{C}\}$ ($S=1, F=0, k_1=n$), then $\mu_\Delta(M) = \rho(M)$.
- If $\Delta = \mathbb{C}^{n \times n}$ ($S=0, F=1, k_1=n$), then $\mu_\Delta(M) = \bar{\sigma}(M)$

Obviously, for a general Δ as in (3.7) we must have

$$\{\delta I_n : \delta \in \mathbb{C}\} \subset \Delta \subset \mathbb{C}^{n \times n}. \quad (3.10)$$

Hence directly from the definition we conclude that

$$\rho(M) \leq \mu_\Delta(M) \leq \bar{\sigma}(M) \quad (3.11)$$

These bounds alone are not sufficient for our purposes, because the gap between ρ and $\bar{\sigma}$ can be arbitrarily large. They are refined by considering transformations on M that do not affect $\mu_\Delta(M)$, but do affect ρ and $\bar{\sigma}$. To do this, define the following two subsets of $\mathbb{C}^{n \times n}$

$$\mathbf{Q} = \{Q \in \Delta : Q^*Q = I_n\} \quad (3.12)$$

$$\mathbf{D} = \left\{ \begin{array}{l} \text{diag}[D_1, \dots, D_S, d_1 I_{m_1}, \dots, d_{F-1} I_{m_{F-1}}, I_{m_F}] : \\ D_i \in \mathbb{C}^{r_i \times r_i}, D_i = D_i^* > 0, d_j \in \mathbb{R}, d_j > 0 \end{array} \right\} \quad (3.13)$$

Note that for any $\Delta \in \Delta, Q \in \mathbf{Q}$, and $D \in \mathbf{D}$,

$$Q^* \in \mathbf{Q} \quad Q\Delta \in \Delta \quad \Delta Q \in \Delta \quad \bar{\sigma}(Q\Delta) = \bar{\sigma}(\Delta Q) = \bar{\sigma}(\Delta) \quad (3.14)$$

$$D\Delta = \Delta D \quad (3.15)$$

Consequently,

Theorem 3.2 For all $Q \in \mathbf{Q}$ and $D \in \mathbf{D}$

$$\mu_\Delta(MQ) = \mu_\Delta(QM) = \mu_\Delta(M) = \mu_\Delta(DMD^{-1}) \quad (3.16)$$

Therefore, the bounds in (3.11) can be tightened to

$$\max_{Q \in \mathbf{Q}} \rho(QM) \leq \mu_\Delta(M) \leq \inf_{D \in \mathbf{D}} \bar{\sigma}(DMD^{-1}) \quad (3.17)$$

The lower bound is always an equality ([Doy]), but $\rho(QM)$ can have multiple local maxima which are not global. Thus local search cannot be guaranteed to obtain μ , but can only yield a lower bound. For computation purposes one can derive a slightly different formulation of the lower bound as a power algorithm which is reminiscent of power algorithms for eigenvalues and singular values ([PacFD]). While there are open questions about convergence, the algorithm usually works quite well and has proven to be an effective method to compute μ .

The upper bound can be reformulated as a convex LMI problem, so the global minimum can, in principle, be found. Unfortunately,

the upper bound is not always equal to μ . For block structures Δ satisfying $2S + F \leq 3$, the upper bound is always equal to $\mu_\Delta(M)$, and for block structures with $2S + F > 3$, there exist matrices for which μ is less than the infimum. The above bounds are much more than just computational schemes. They are also theoretically rich, and can unify a number of apparently quite different results in linear systems theory, as will be seen below and in the companion papers.

3.2 μ and LFTs

Consider a complex matrix M partitioned as in (2.1) and suppose there are two defined block structures, Δ_1 and Δ_2 , which are compatible in size with M_{11} and M_{22} respectively. Define a third structure $\Delta = \{\text{diag}(\Delta_1, \Delta_2) : \Delta_1 \in \Delta_1, \Delta_2 \in \Delta_2\}$. Now there are three structures with which we may compute μ with respect to. The notation we use to keep track of this is as follows: $\mu_1(\cdot)$ is with respect to Δ_1 , $\mu_2(\cdot)$ is with respect to Δ_2 , $\mu_\Delta(\cdot)$ is with respect to Δ . In view of this, $\mu_1(M_{11})$, $\mu_2(M_{22})$ and $\mu_\Delta(M)$ all make sense, though for instance, $\mu_1(M)$ does not. The constant matrix problem we wish to solve is:

- determine whether the LFT $F_l(M, \Delta_2)$ is well posed for all $\Delta_2 \in \Delta_2$ with $\bar{\sigma}(\Delta_2) \leq \beta$, and,
- if so, then determine how “large” $F_l(M, \Delta_2)$ can get for this norm-bounded set of perturbations.

The following three simple theorems answer this problem. The first theorem is nothing more than a restatement of the definition of μ .

Theorem 3.3 The linear fractional transformation $F_l(M, \Delta_2)$ is well posed for all $\Delta_2 \in \mathbf{B}\Delta_2$ if and only if $\mu_2(M_{22}) < 1$.

As the “perturbation” Δ_2 deviates from zero, the matrix $F_l(M, \Delta_2)$ deviates from M_{11} . The range of values that $\mu_1(F_l(M, \Delta_2))$ takes on is intimately related to $\mu_\Delta(M)$, as follows:

Theorem 3.4 (MAIN LOOP THEOREM) The following are equivalent:

$$\mu_\Delta(M) < 1 \iff \begin{cases} \mu_2(M_{22}) < 1, \text{ and} \\ \max_{\Delta_2 \in \mathbf{B}\Delta_2} \mu_1(F_l(M, \Delta_2)) < 1 \end{cases}$$

This theorem forms the basis for all uses of μ in linear system robustness analysis, whether from a state-space, frequency domain, or Lyapunov approach. The importance of the theorem is summarized as: Suppose a property \mathcal{P} , of a matrix W can be related to a “ μ test” on the matrix, in particular,

$$\text{matrix } W \text{ satisfies property } \mathcal{P} \iff \mu_{\Delta_p}(W) < 1.$$

Then the linear fractional transformation $F_l(M, \Delta)$ is well defined, and has the property \mathcal{P} for every $\Delta \in \mathbf{B}\Delta$ if and only if $\mu_{\tilde{\Delta}}(M) < 1$, where $\tilde{\Delta} := \{\text{diag}[\Delta_p, \Delta] : \Delta_p \in \Delta_p, \Delta \in \Delta\}$. In other words, whenever a property of a matrix can be related to a μ test, then there will be a μ test (of greater complexity) to determine if the property is robust to structured linear fractional transformations.

The role of the block structure Δ_2 in the MAIN LOOP theorem is clear - it is the structure that the perturbations come from. However the role of the perturbation structure Δ_1 is often misunderstood. Note that $\mu_1(\cdot)$ appears on the right hand side of the theorem, so that the set Δ_1 defines what particular property of $F_l(M, \Delta_2)$ is considered.

As an example, consider the following version of the small gain theorem using the block structures $\Delta_1 := \{\delta_1 I_n : \delta_1 \in \mathbb{C}\}$, $\Delta_2 = \mathbb{C}^{m \times m}$, and

$$\Delta := \left\{ \begin{bmatrix} \delta_1 I_n & 0_{n \times m} \\ 0_{m \times n} & \Delta_2 \end{bmatrix} : \delta_1 \in \mathbb{C}, \Delta_2 \in \mathbb{C}^{m \times m} \right\} \subset \mathbb{C}^{(n+m) \times (n+m)}$$

with M from the state space model in (2.4). Hence $\mu_1(A) = \rho(A)$ and $\mu_2(D) = \bar{\sigma}(D)$. Applying the theorem with this data gives that the following are equivalent:

$$\begin{aligned} & \rho(A) < 1 \quad \text{and} \quad \max_{\substack{\delta_1 \in \mathbb{C} \\ |\delta_1| < 1}} \bar{\sigma}(D + C\delta_1(I - A\delta_1)^{-1}B) < 1 \\ \text{iff} \quad & \bar{\sigma}(D) < 1 \quad \text{and} \quad \max_{\substack{\Delta_2 \in \mathbb{C}^{m \times m} \\ \bar{\sigma}(\Delta_2) \leq 1}} \rho(A + B\Delta_2(I - D\Delta_2)^{-1}C) < 1 \\ \text{iff} \quad & \mu_\Delta(M) < 1. \end{aligned}$$

Note that the $\max \bar{\sigma}$ is just the H_∞ norm of the transfer function.

3.3 State-Space/Frequency Domain Tests

It is known [PacD], that when the block structure includes repeated scalars ($S > 0$), it is possible that the upper bound is much larger than μ . How should these differences be interpreted when using the upper bound to perform analysis tests for robustness properties? We answer that question by considering the robust performance problem for an uncertain difference equation, and outline the various tests and corresponding conclusions (see also [BoyY] and [PacD3] for additional interpretations). Performance will be characterized in terms of the induced l_2 gain from disturbance to error, which for LTI systems, is the same as the $\|\cdot\|_\infty$ norm of the transfer function.

Consider the uncertain difference equation in (2.5) and (2.6). Define three augmented block structures, Δ_N , Δ_S and Δ_P as

$$\Delta_N := \{\text{diag}[\delta_1 I_n, \Delta_2] : \delta_1 \in \mathbb{C}, \Delta_2 \in \mathbb{C}^{n_p \times n_p}\} \quad (3.18)$$

$$\Delta_S := \{\text{diag}[\Delta_N, \Delta] : \Delta_N \in \Delta_N, \Delta \in \Delta\} \quad (3.19)$$

$$\Delta_P := \{\text{diag}[\Delta_2, \Delta] : \Delta_2 \in \mathbb{C}^{n_p \times n_p}, \Delta \in \Delta\} \quad (3.20)$$

along with the corresponding scaling sets \mathbf{D}_N , \mathbf{D}_S and \mathbf{D}_P . We begin with the main result for linear, time-invariant perturbations, [DoyWS], [DoyP].

Theorem 3.5 (Time-invariant, robust performance) *Given the matrices and sets as defined above, the following conditions are equivalent:*

1. *There exists a constant $\beta \in [0, 1)$ such that for each fixed $\Delta \in \mathbf{B}\Delta$, the uncertain system (2.6) is well-posed ($I - M_{33}\Delta$ is invertible), stable, and for zero-initial-state-response, the error e satisfies $\|e\|_2 \leq \beta \|d\|_2$*
2. $\mu_\Delta(M_{33}) < 1$ and $\max_{\Delta \in \mathbf{B}\Delta} \mu_{\Delta_N}(F_1(M, \Delta)) < 1$
3. $\mu_{\Delta_S}(M) < 1$ (SS μ)
4. $\rho(M_{11}) < 1$ and $\max_{\theta \in [0, 2\pi]} \mu_{\Delta_P}(F_u(M, e^{j\theta} I_n)) < 1$ (FD μ).

Item 1 in this theorem is the desired robust performance conclusion. Item 2 rephrases Item 1, using the μ characterization of $\|\cdot\|_\infty < 1$. Items 3 and 4, known respectively as the “state-space μ test (SS μ)” and the “frequency domain μ test (FD μ)” are the computational tests involving computing μ for various matrices. We will investigate the additional conclusions that are possible when the $\bar{\sigma}(DMD^{-1})$ upper bound is used to implement the computational tests of items 3 and 4. The FD μ test is the most common use of μ , essentially Bode plots with $\mu(\cdot)$ replacing $\bar{\sigma}(\cdot)$. The SS μ test was introduced in [DoyP].

Using the $\bar{\sigma}(DMD^{-1})$ upper bound in place of μ , sufficient conditions for robust performance are the “state-space upper bound test” and the “frequency domain upper bound test”.

- 3' $\inf_{D_S \in \mathbf{D}_S} \bar{\sigma}(D_S M D_S^{-1}) < 1$ (SSUB)
- 4' $\max_{\theta \in [0, 2\pi]} \inf_{D_P \in \mathbf{D}_P} \bar{\sigma}[D_P F_u(M, e^{j\theta} I_n) D_P^{-1}] < 1$ (FDUB)

While the various μ tests given in Theorem 3.5 are all equivalent, these two upper bound tests are **very different**. In particular, recalling that the upper bound is equal to μ when $S = F = 1$, the SSUB condition is actually equivalent to

$$\inf_{D_P \in \mathbf{D}_P} \max_{\theta \in [0, 2\pi]} \bar{\sigma}[D_P F_u(M, e^{j\theta} I_n) D_P^{-1}] < 1. \quad (3.21)$$

This condition is much stronger than the frequency domain upper bound test, since in (3.21), the same $D_P \in \mathbf{D}_P$ must work for all $\theta \in [0, 2\pi]$. For that reason, we call equation (3.21) the *frequency domain constant D test, FDCD*. Now, from strongest to weakest, the various conditions are:

$$\begin{aligned} & \inf_{D_S \in \mathbf{D}_S} \bar{\sigma}(D_S M D_S^{-1}) < 1 && \text{(SSUB)} \\ & \Downarrow \\ & \inf_{D_P \in \mathbf{D}_P} \max_{\theta \in [0, 2\pi]} \bar{\sigma}[D_P F_u(M, e^{j\theta} I_n) D_P^{-1}] < 1 && \text{(FDCD)} \\ & \Downarrow && \text{gap} \\ & \max_{\theta \in [0, 2\pi]} \inf_{D_P \in \mathbf{D}_P} \bar{\sigma}[D_P F_u(M, e^{j\theta} I_n) D_P^{-1}] < 1 && \text{(FDUB)} \\ & \Downarrow && \text{gap} \\ & \max_{\theta \in [0, 2\pi]} \mu_{\Delta_P}(F_u(M, e^{j\theta} I_n)) < 1 && \text{(FD}\mu\text{)} \\ & \Downarrow \\ & \mu_{\Delta_S}(M) < 1 && \text{(SS}\mu\text{)} \\ & \Downarrow \end{aligned}$$

ROBUST PERFORMANCE with LTI perturbations

Note in both instances where the implication is given as \Downarrow rather than \Downarrow , there truly is a gap. Also, there are two such “gaps” between the state space tests, SSUB and SS μ , while there is only one gap between the frequency domain tests, FDUB and FD μ . The top conditions are the strongest, and are equivalent to a very strong form of robust Lyapunov stability, [BoyY]. The SSUB and SS μ tests can be thought of as *defining* two extreme notions of robust stability and performance. The motivation for this will be discussed in later, where we will refer to the SSUB test as \mathcal{Q} stability and performance and the SS μ test as μ stability and performance.

Given that the upper bound is computable, which test should be used, the state space upper bound test, SSUB (equivalently FDCD), or the frequency domain upper bound test, FDUB? The answer depends on the assumptions that are made about the perturbations.

State space upper bound test

If the SSUB is used, and successful, then the robust performance conclusion holds for time-varying perturbations (and with proper interpretation, cone bounded nonlinear perturbations).

Theorem 3.6 *Let M be given as in (2.5), along with an uncertainty structure Δ . If there is a $D_S \in \mathbf{D}_S$ such that*

$$\bar{\sigma}(D_S M D_S^{-1}) = \beta < 1 \quad (3.22)$$

then there exist constants $c_1 \geq c_2 > 0$, such that for all perturbation sequences $\{\Delta_k\}_{k=0}^\infty$ with $\Delta_k \in \Delta$, $\bar{\sigma}(\Delta_k) < \frac{1}{\beta}$, the time-varying, uncertain system

$$\begin{bmatrix} x_{k+1} \\ e_k \end{bmatrix} = F_1(M, \Delta_k) \begin{bmatrix} x_k \\ d_k \end{bmatrix} \quad (3.23)$$

is zero-input, exponentially stable, and furthermore, if $\{d_k\}_{k=0}^\infty \in l_2$, then

$$c_2(1 - \beta^2) \|x\|_2^2 + \|e\|_2^2 \leq \beta^2 \|d\|_2^2 + c_1 \|x_0\|^2$$

In particular, $\|e\|_2^2 \leq \beta^2 \|d\|_2^2 + c_1 \|x_0\|^2$.

Remark: Note that D_S will appear as $D_S = \text{diag}[D_1, d_2 I, D]$, where $D_1 = D_1^* > 0, D_1 \in \mathbb{C}^{n \times n}$. The constants c_1 and c_2 are the maximum and minimum singular values of D_1 .

Frequency domain upper bound test

Since FDUB is a weaker condition than the SSUB, it is “closer” to the exact condition for robust performance under LTI perturbations. Therefore, if the perturbations are better modeled as linear, time-invariant perturbations, this frequency domain test is more appropriate, but no general conclusion can be reached about time-varying perturbations. Nevertheless, some results using frequency dependent scalings and non-LTI perturbations are indeed possible. In [Saf] multiplier methods are used to derive restrictions on the frequency dependent scalings $D(\theta)$, which yield robustness theorems for slope-bounded, nonlinear perturbations. In [PacT], frequency dependent scalings are used with the swapping lemma to get robustness bounds on time-varying perturbations, in the form of absolute value bounds on the perturbations and their time-derivatives.

4 μ , Q , and L_1

We have considered several different measures of robust stability and performance for the system in (2.6), from $SS\mu$ to the SSUB. We will concentrate on these two measures, and compare them briefly with another very important measure that has emerged in the L_1 theory of robust performance with structured uncertainty. Space constraints preclude a review of the L_1 theory, which has undergone a dramatic and impressive development in the last 5 years in the work of Dahleh, Khamash, Ohta, Pearson, and coauthors (see [KhaP] and references therein). For simplicity, we will refer to the $SS\mu$ test as simply μ and the SSUB upper bound as Q , since it is directly related to quadratic stability, and focus our attention on the robust performance problem, which clearly includes robust stability as a special case.

The μ , Q , and L_1 tests all guarantee robust performance, but with different assumptions about perturbations and the norm for the performance objective. The μ and Q theories are for L_2 induced norms, while the L_1 theory is for L_∞ induced norms. The other distinction is that the μ theory treats LTI perturbations, and the Q and L_1 handle Nonlinear and Time-Varying perturbations (NTV). This is summarized in the table below.

	LTI	NTV
L_2	μ	Q^*
L_∞	μ	L_1

Note that with the exception of Q each test is necessary and sufficient for robust performance, and it is conjectured that Q is necessary and sufficient for certain block structures. Recall that in general, μ is computed using bounds, but that Q involves solving LMIs, so is attractive computationally. Interestingly, L_1 is probably the easiest to compute, involving only the computation of L_1 norms and finding the spectral radius of a positive matrix ([KhaP]).

As a final comparison, it can be easily shown that the tests are ordered, with

$$\mu \leq Q \leq L_1 \quad (4.24)$$

What this means is that for a given system, if the Q test passes, the μ test must, and similarly for L_1 and Q . That $\mu \leq Q$ was shown above, and $Q \leq L_1$ follows immediately from the equivalence of the SSUB and the FDCD problems and the fact that the L_1 norm of a convolution kernel is greater than the H_∞ norm of its transform, together with the results in [KhaP]. The inequalities are typically strict and it is possible for the gaps to be arbitrarily large.

It is not clear exactly what are the implications of these results for control design as well as for further research. Clearly there is a need for more refined results, the ability to combine LTI and NTV uncertainty and exploit additional structure such as the slowly-varying nature of some perturbations. The results in [Saf] and [PacT] suggest how this might be done in the LFT/ μ/Q framework, but much more work is

needed. We also need much more precise modeling and ID methods to exploit the detailed structure of the uncertainty in our models.

If one accepts Q as the measure of robust performance, a rich theory can be developed, with generalizations to uncertain systems of conventional theories of robust stability and performance, balanced realizations, model reduction, stabilization, optimal control, and model validation. These will be pursued in the companion papers. It is not surprising that the easiest generalizations of standard results to uncertain LFT systems is done using the Q framework. Indeed, most of the standard results rely on Q machinery, but since μ and Q are the same for these simple block structures, we are less aware of the distinction. Once we begin extending our results to systems with uncertainty, the distinction becomes significant. Of course, a key feature of the Q theory is that computation involves solving LMIs.

5 Linear Matrix Inequalities (LMIs)

The general LMI problem involves sets of the form

$$\mathcal{X} = \left\{ \begin{array}{l} \text{diag}[X_1, \dots, X_S, x_1 I, \dots, x_F I] : \\ X_i \in \mathbf{C}^{r_i \times r_i}, X_i = X_i^*, x_j \in \mathbf{R} \end{array} \right\} \quad (5.25)$$

and a list of matrices A_i, B_i, C_i, D_i . The simplest general LMI problem is whether there exists $X \in \mathcal{X}$ such that

$$A_i^* X A_i - B_i^* X B_i + X C_i + C_i^* X + D_i < 0 \quad \forall i$$

Depending on the particular problem, the $<$ may be a \leq . It is easy to see that such conditions produce a set of solutions which are convex, which makes LMIs attractive computationally. This is a decision problem; the answer is yes or no. Sometimes, however, the A_i, B_i, C_i , and D_i are functions of a real, positive parameter α , and we want to know, for example, what is the largest α for which there is no solution. Typically this involves an iteration on α , and consequently, answering the decision question many times.

Numerical methods for the solution of LMI problems are reviewed in [Beck]. In this section we give several standard examples of LMIs. The companion papers to this paper consider further applications of LMIs.

Stability

A discrete-time system is stable iff $\rho(A) < 1$. It is well-known that this is equivalent to the following conditions.

$$\begin{aligned} & \rho(A) < 1 \\ \iff & \exists T : \bar{\sigma}(TAT^{-1}) < 1 \\ \iff & \exists T : TAT^{-1}(TAT^{-1})^* - I < 0 \\ \iff & \exists T : A(T^*T)^{-1}A^* - (T^*T)^{-1} < 0 \\ \iff & \exists X > 0 : AXA^* - X < 0 \end{aligned} \quad (5.26)$$

Note that the last condition involves 2 LMIs on an unstructured X .

Similarly, a continuous time system is stable iff all eigenvalues of A are in the open left half plane, or equivalently,

$$\exists X > 0 : AX + XA^* < 0$$

Stabilization

Consider a discrete time system with realization $G(z) = \begin{bmatrix} A & B \\ I & 0 \end{bmatrix}$ with $A \in \mathbf{R}^{n \times n}$ and $B \in \mathbf{R}^{n \times p}$.

The following lemma characterizes the stabilizability of G , or equivalently, the stabilizability of the pair (A, B) in terms of 2 LMIs.

Lemma 5.1 *Suppose the matrix B in $G(z)$ is of full column rank (with $\text{rank}(B) = p < n$), where $B_\perp \in \mathbf{R}^{n \times (n-p)}$ is such that $B_\perp^* B = 0$ and $\begin{bmatrix} B & B_\perp \end{bmatrix}$ is invertible. Then there exists a static feedback matrix F*

such that the closed loop system matrix $A + BF$ is stable if and only if there exist a $X > 0$ such that

$$B_1^* A X A^* B_1 - B_1^* X B_1 < 0 \quad (5.27)$$

μ Upper Bound

The upper bound from (3.17) may be expressed as an LMI.

$$\begin{aligned} \mu(M) < \beta & \\ \Leftrightarrow \exists \hat{D} > 0 : \bar{\sigma}(\hat{D}M\hat{D}^{-1}) < \beta & \\ \Leftrightarrow \exists \hat{D} > 0 : \hat{D}^{-1}M^*\hat{D}^2M\hat{D}^{-1} - \beta^2 I < 0 & \\ \Leftrightarrow \exists \hat{D} > 0 : M^*\hat{D}^2M - \beta^2\hat{D}^2 < 0 & \\ \Leftrightarrow \exists X > 0 : M^*XM - \beta^2X < 0 & \end{aligned} \quad (5.28)$$

Note that (5.26) is a special case of (5.28).

H_∞ Norms and Riccati Equations

Let $M(z) = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \in \mathcal{RH}_\infty$ and define

$$\begin{aligned} E &:= A + B(I - D'D)^{-1}D'C \\ G &:= -B(I - D'D)^{-1}B' \\ Q &:= C'(I - DD')^{-1}C \end{aligned}$$

Suppose E is nonsingular and define a symplectic matrix as

$$S := \begin{bmatrix} E + GE'^{-1}Q & -GE'^{-1} \\ -E'^{-1}Q & E'^{-1} \end{bmatrix}$$

Then the following statements are equivalent:

- (a) $\|M(z)\|_\infty < 1$
- (b) S has no eigenvalues on the unit circle and $\|C(I - A)^{-1}B + D\| < 1$
- (c) $\exists X \geq 0$ with $I - D'D - B'XB > 0$, $(I + GX)^{-1}E$ stable, and

$$E'XE - X - E'XG(I + XG)^{-1}XE + Q = 0$$

- (f) $\exists X > 0$ such that $I - D'D - B'XB > 0$ and

$$E'XE - X - E'XG(I + XG)^{-1}XE + Q < 0$$

- (g) $\exists X > 0$ such that

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}' \begin{bmatrix} X & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} A & B \\ C & D \end{bmatrix} - \begin{bmatrix} X & 0 \\ 0 & I \end{bmatrix} < 0$$

- (h) $\exists T$ nonsingular such that

$$\bar{\sigma} \left(\begin{bmatrix} TAT^{-1} & TB \\ CT^{-1} & D \end{bmatrix} \right) = \bar{\sigma} \left(\begin{bmatrix} T & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} T & 0 \\ 0 & I \end{bmatrix}^{-1} \right) < 1$$

Note that (h) is the SSUB and is equal to μ because of the block structure. It is equivalent to (g), which is 2 LMIs. The connection between (f) and (g) is just the Schur complement formula for positive definite matrices.

6 Towards a postmodern control theory

This session is on the theoretical end of what might be called postmodern control theory. While it may seem pretentious to use such a term, there is emerging a philosophy and set of techniques that seem distinct enough from classical and modern control to warrant a new label. Without getting into the philosophy here, it may be useful to outline how the methods of postmodern control compare with their classical and modern origins.

The central object of study in classical control, the transfer function, was displaced in modern control by state-space theory. In postmodern control, the central objects of study are LFTs, and their generalizations to nonlinear systems. While classical control was concerned with issues like margins and bandwidth, modern control focused on

such things as eigenvalues, norms, and variances. Singular values and MIMO robust stability results provided a starting point for a blend of the classical and modern, and a bridge to the postmodern, with its focus on robust performance with structured uncertainty using μ , Q , and L_1 theories.

Classical loopshaping in its various guises gave way to optimal control in the modern era. A bridge was formed by multivariable loopshaping and H_∞ , but a fully developed synthesis theory for robust performance has yet to be developed, in spite of the recent success of μ -synthesis. The results in [PZPB] suggest what directions a postmodern synthesis theory might take.

Throughout this session, an emphasis is placed on reducing a variety of problems to solving LMIs. LMIs play the same central role in the postmodern theory as Lyapunov and Riccati equations played in the modern, and in turn various graphical techniques such as Bode, Nyquist and Nichols plots played in the classical.

The companion papers use the postmodern machinery reviewed here to extend standard modern results in a number of directions. In [WDBG] balanced truncation model reduction is extended to uncertain LFT systems, with similar extensions of the parametrization of all stabilizing controllers in [LuZD] and H_∞ theory in [PZPB]. The LFT machinery not only extends the standard results in important ways, it simplifies the proofs, often substantially. Exciting new developments in handling real parametric uncertainty are reviewed in [YoND], while recent work on model validation is presented in [Newl]. In all cases, LMIs play a central role in the computation.

7 References

- Beck: C. Beck, "Computational Issues in Solving LMIs," 1991 IEEE CDC.
- BoyY: S. Boyd, and Q. Yang, "Structured and simultaneous Lyapunov functions for system stability problems," *IJC*, 1989, vol. 49, No. 6, 2215-2240.
- Doy1: J.C. Doyle, "Analysis of feedback systems with structured uncertainties," *IEE Proceedings*, vol. 129, Part D, No. 6, pp. 242-250, Nov. 1982.
- Doy2: J.C. Doyle, "Structured uncertainty in control system design", IEEE CDC, December 1985, Ft. Lauderdale.
- DoyP: J.C. Doyle and A. Packard, "Uncertain multivariable systems from a state space perspective", 1987 ACC, pp. 2147-2152.
- DoyWS: J.C. Doyle, J. Wall and G. Stein, "Performance and robustness analysis for structured uncertainty," in 1982 IEEE CDC, pp. 629-636.
- FanT: M.K.H. Fan and A.L. Tits, "Characterization and Efficient Computation of the Structured Singular Value", *IEEE TAC*, vol. AC-31, no. 8, pp. 734-743, Aug. 1986.
- KhaP: M. Khammash and J. B. Pearson, "Performance Robustness of Discrete-Time Systems with Structured Uncertainty", *IEEE TAC*, vol. AC-36, no. 4, pp. 398-412, 1991.
- LuZD: W. Lu, K. Zhou, and J. Doyle, "Stability and stabilization of LSI multidimensional systems," 1991 IEEE CDC.
- MorM: B. Morton and R. McAfoos, "A Mu-Test for Real-Parameter Variations," 1985 ACC, Boston.
- Newl: M. Newlin, "Model validation and generalized μ " 1991 IEEE CDC.
- PacD1: Andy Packard and John Doyle, "Structured singular value with repeated scalar blocks," 1988 ACC, pp. 1213-1218, Atlanta.
- PacD2: Andy Packard and John Doyle, "Robust Control of Multivariable and Large Scale Systems," Final Technical Report, Contract No. F49620-86-C-0001, March 1988, prepared for AFOSR.
- PacD3: Andy Packard and John Doyle, "Quadratic stability with real and complex perturbations," *IEEE TAC*, vol. AC-35, no. 1, pp. 25-38, Feb. 1990.
- PacFD: A. Packard, M. Fan, and J. Doyle, "A power method for the structured singular value," 1988 IEEE CDC, pp. 2132-2137, Austin.
- PacT: A. Packard and J. Teng, "Robust stability with time-varying perturbations," 28th Allerton Conference, October, 1990, pp. 765-767.
- PZPB: A. Packard, K. Zhou, P. Pandey, G. Becker, "A collection of robust control problems leading to LMIs," 1991 IEEE CDC.
- Red: R. Redheffer, "Inequalities for a matrix Riccati equation," *Journal of Mathematics and Mechanics*, vol. 8, no. 3, 1959.
- Saf: M.G. Safonov, "Stability of interconnected systems having slope bounded nonlinearities", 8th International Conf. on Analysis and optimization of systems, Nice, France, June 19-22, 1984.
- WDBG: W. Wang, J.C. Doyle, C. Beck and K. Glover, "Stability and model reduction of linear fractional transformations" 1991 IEEE CDC.
- YoND: P. Young, M. Newlin, and J. Doyle, "Structured Singular Value Analysis with real and complex uncertainties," 1991 IEEE CDC.